The Pennsylvania Department of Education

Proposal to the US Department of Education for Participation in the No Child Left Behind (NCLB) Growth Model Pilot Program

October 31, 2006

Submitted by:

Gerald L. Zahorchak, D.Ed.
Secretary of Education
Commonwealth of Pennsylvania

Table of Contents

INTRODUCTION **3**

Proposed Accountability Plan Amendment 4
Pennsylvania's Current Accountability Workbook: Status Measures 4
Pennsylvania's Current Accountability Workbook: Improvement Measures 4
Pennsylvania's Proposed Accountability Workbook Amendment: Growth Measures 4
Proposed Pennsylvania Growth Model as an Additional AYP Method 5
Scope of Proposed Pennsylvania Pilot 7
High Standards in Pennsylvania 7

PENNSYLVANIA GROWTH MODEL AMENDMENT SCHEME **8**

Timeline for Implementation of Proposed Amendment 9
Growth Model Integrity 9
Conclusion 9

RESPONSE TO USED SEVEN CORE PRINCIPLES **10**

Core Principle 1: 100% Proficiency by 2014 and Incorporating Decisions about Student Growth into School Accountability 10
Core Principle 2: Establishing Appropriate Growth Targets at the Student Level 19
Core Principle 3: Accountability for Reading/Language Arts and Mathematics Separately 22
Core Principle 4: Inclusion of All Students 23
Core Principle 5: State Assessment System and Methodology 27
Core Principle 6: Tracking Student Progress 33
Core Principle 7: Participation Rates and Additional Academic Indicator 36

APPENDIX A - HISTORY OF GROWTH MODELS IN PENNSYLVANIA **38**

APPENDIX B - INTEGRATION OF GROWTH MODEL **40**

APPENDIX C – REFERENCES **41**

APPENDIX D – MEMBERS OF THE PENNSYLVANIA VALUE-ADDED ASSESSMENT SYSTEM STATEWIDE WORK GROUP **42**

APPENDIX E – DEFINITIONS OF AYP METHODS **43**

APPENDIX F - PROJECTION METHODOLOGY **44**

APPENDIX G – MATCH RATES FOR 2006 GRADES 4 AND 6 ..... **46**

Introduction

The No Child Left Behind Act (NCLB) has clearly stressed the expectation that all students must be proficient in reading and math by the year 2014. As part of this expectation, each state has been given the task of developing an accountability workbook that outlines how it will monitor progress toward this goal. Recently, the U.S. Department of Education offered states the opportunity to submit a proposal to participate in a pilot program that uses measures of student longitudinal growth as part of the adequate yearly progress (AYP) calculations. These growth models must not delay the expectations of NCLB. Instead, states were invited to propose metrics that would provide an additional way to identify and recognize schools that had placed all of their students on a trajectory to proficiency.

The Commonwealth of Pennsylvania and the Pennsylvania Department of Education (PDE) currently provide two growth models for school districts. The first, a value-added model, used for school improvement purposes, allows schools to compare themselves to a growth standard, and provides valuable insights regarding whether or not their students *as a group* made one year's worth of progress. In contrast, the second model uses a projection methodology that allows schools to determine if individual students are on a trajectory to achieve and maintain proficiency according to state standards. Pennsylvania believes that the *projection to proficiency* methodology, with its emphasis on individual student growth, offers great promise as a potential addition to the cadre of tools available to monitor movement toward NCLB goals. Therefore, the Pennsylvania Department of Education proposes that an amendment to Pennsylvania's present Accountability Workbook be accepted for the determination of Adequate Yearly Progress (AYP) for students enrolled in Pennsylvania public schools during academic year 2006-07.

The proposed amendment specifies the inclusion of a projection to proficiency/growth metric – NOT a value-added metric – to assess the effectiveness of including a longitudinal analysis of student achievement data for determination of AYP status. The inclusion of this amendment will recognize schools in which students have not yet achieved proficiency but have demonstrated significant growth towards proficiency in a time frame aligned to Pennsylvania's Annual Measurable Objective (AMO) targets. It will inform Pennsylvania schools that they should address the growth of students who are presently rated as proficient so that these students will continue to perform in the proficient range in the future. The Pennsylvania Department of Education strongly believes that its proposed projection to proficiency growth model supports the goals of NCLB and that it will encourage school districts to place at-risk students on an accelerated path to proficiency in both reading and math through the targeted use of resources, interventions, professional development and high standards.

Proposed Accountability Plan Amendment

Pennsylvania proposes to apply an individual student "projection to proficiency" metric as a method for schools and districts to meet AYP in addition to the methods presently defined and accepted in the USED approved Pennsylvania Accountability Workbook.

Pennsylvania's Current Accountability Workbook: Status Measures

Pennsylvania's current Accountability Workbook outlines the process to measure schools' AYP through status targets that increase toward the No Child Left Behind (NCLB) goal of 100% proficient by 2014. Hill, Gong, Marion, DePascale, Dunn, and Simpson (2005) define status models as the performance of a school at any given time. Currently, Pennsylvania uses student performance on the PSSA in Grades 3 through 8 and 11 in reading and math as the "status indicator" for NCLB proficiency targets. AYP is accomplished by achieving designated yearly NCLB performance targets in reading and math at the district, school, and subgroup levels, and graduation and participation rates. Confidence intervals are applied to these measures to account for any statistical error in Pennsylvania's metric.

Pennsylvania's Current Accountability Workbook: Improvement Measures

Pennsylvania also includes "improvement indicators" in its Accountability Workbook. These improvement measures look at status relative to the prior status of a previous cohort of students at the same grade level (Hill et al., 2005) and include Safe Harbor calculations and the Pennsylvania Performance Index (PPI). Page eight of the proposal offers a visual representation of the current status and improvement measures in use. If schools do not achieve AYP status targets, they may use improvement measures in Pennsylvania's current Accountability Workbook.

Pennsylvania's Proposed Accountability Workbook Amendment: Growth Measures

While many of Pennsylvania's schools have been able to achieve the designated AYP targets through the current status and improvement measures, there remain a number of schools that are helping their students grow toward proficiency, but are not recognized for this accomplishment. Therefore, Pennsylvania is proposing the use of a "growth measure," specifically the projection capability of the Pennsylvania Value-Added Assessment System (PVAAS), as another option for schools to meet AYP. A true "growth measure" evaluates the status of a particular group of students against the prior performance of these same students (Hill et al., 2005). By recognizing the importance of accelerated student progress to proficiency through the use of a growth measure, Pennsylvania seeks to recognize schools that have made progress that is aligned to the proficiency target but whose current level of achievement may not yet reflect the designated benchmarks of the Accountability Workbook.

Proposed Pennsylvania Growth Model as an Additional AYP Method

Under the proposed accountability system, districts, schools and subgroups will have three options for meeting AYP proficiency targets in reading and math: (1) status, (2) safe harbor/improvement, or (3) projection. Pennsylvania proposes to use the ***projection*** metric of PVAAS to estimate the score of a particular student on a future state assessment that will then be used as part of the AYP determination. Using all available achievement data on the students, the projection calculation estimates a student's performance on a future assessment based on the student's test performance history and the histories of students with similar performance patterns. The individual student projection data will be used to determine the percent of students by district, school, subgroup and subject area who are projected to attain proficiency on a future Pennsylvania State System of Assessment (PSSA) examination as specified in Table 1:

Table 1: Present and Projected Grades

Present Grade	4	5	6	7	8
Projected to Proficiency in Grade	6	7	8	8	11

The grades chosen for the projection to proficiency are based on the varied school configurations presently utilized in Pennsylvania.

Demographic data are <u>not</u> used in the analyses. Because the calculation uses longitudinal data from students with similar test performance patterns and uses all available test data in the database of performance measures, the projection model produces reliable estimates of projected performance on future state assessments.

Based on the students' projection calculations, schools will be assigned credit for all students who are projected to be proficient in a specified grade in the future, whether they are currently below proficient or are currently proficient. It does not assign schools any credit for students who are currently proficient but are projected to score below proficient on the future assessment. It does not assign schools any additional credit for students who score at the advanced level. For example, based on the projection calculations, Student A in 4th grade who scored below the 4th grade proficiency standard is projected to be proficient if the projected 6th grade score for that student is greater than or equal to the Grade 6 proficient standard. Similarly, Student B in 4th grade who scored above the 4th Grade proficiency standard would be projected to be non-proficient if the projected 6th grade score for that student is less than the Grade 6 proficient standard. These projected student scores are then used to determine the number of students projected to be proficient at the district, school and subgroup level. The projection model will apply to all students in the school or subgroup of interest regardless of the proficiency determination of the students based on status measures. The percent of students projected to be proficient on a future PSSA examination will then be used to determine AYP status based on the presently accepted Annual Measurable Objectives (AMO) status targets using the following rule:

Districts and schools meet AYP proficiency requirements if *the district and school and all subgroups* meet the

annual measurable objective in *both reading and mathematics* either by meeting the status or improvement measures in the current approved Pennsylvania Accountability Workbook or through the proposed projection model. The proposed projection model will be applied for either mathematics or reading subject areas or both depending on the outcome of the status and improvement measure AYP determination. When using the proposed projection model to determine AYP:

1. Each district, school or subgroup's percentage of students who are projected to score proficient or advanced on reading meets the approved annual measurable objective for reading; or

2. Each district, school, or subgroup's percentage of students who are projected to score proficient or advanced on mathematics meets the approved annual measurable objective for mathematics.

In all cases, projections to performance on a future PSSA examination will be calculated when sufficient data to calculate the projection is available.

Pennsylvania's projection model keeps schools on target of 100% of students reaching proficiency by 2013-14 by incorporating the already approved AYP proficiency targets. The AMO's have been approved in Pennsylvania's Accountability Workbook and increase over time until they reach the goal of 100% proficiency in 2013-14. Pennsylvania further proposes that the amendment include consideration of a 68% confidence interval for the percent of students projected to be proficient on a future PSSA examination because of the statistical error that is included in any metric. Confidence intervals are used to control for sampling errors or measurement errors, promote fairness and equity, and increase the validity and reliability of the accountability decisions. In Status models, confidence intervals help to correct for sampling error (the student members of a school are fluid from year to year and therefore, students in a given year are considered a sample not a population), to reduce the risk of Type 1 error (classifying a school as needing improvement when in fact it does not) and ideally, to help correct for measurement error in the assessments. In a growth model, the use of confidence intervals follows a similar rationale; they are needed to correct for sampling error, Type 1 error and measurement error. It is still the case with this growth model that confidence intervals are not only appropriate but also necessary to reduce sampling error. There is sampling error with growth models since the results are generalized from the students tested to all students who attend or will attend the school. In other words, the intent of any accountability system under NCLB, whether it is status or growth, is to use information gathered from one set of students tested to make determinations about the school as a whole. Generalizations from one cohort to another are subject to sampling error. In addition, because the growth model scores are based on score projections, which have errors marginally higher than observed scores, the use of confidence intervals to correct for measurement error is even more critical.

This request is consistent with the utilization of confidence intervals in the current, approved Pennsylvania Accountability Workbook for status and safe harbor calculations. Pennsylvania's request for approval of a 68%

confidence interval based on the measure of one standard error of the projection estimate is a conservative but necessary interval approximation of the true percent of students projected to proficiency. Pennsylvania further requests that the inclusion of a confidence interval for the percent proficient as defined by projection to proficiency be considered as a separate corollary to its main proposal as outlined above.

Scope of Proposed Pennsylvania Pilot

Pennsylvania has implemented a Growth Model since School Year (SY) 2001-2002 involving 100 school districts initially, with statewide implementation including all 501 school districts in SY05-06. The Commonwealth intends to implement its amendment, if approved, with all 501 school districts for AYP determinations in the current school year (SY06-07).

High Standards in Pennsylvania

By incorporating the proposed growth model into AYP calculations, Pennsylvania is creating an opportunity to reinforce the goals of NCLB by acknowledging schools that are heading in the direction of meeting AYP targets, while still maintaining high expectations for all students. The amended accountability system will encourage schools and districts to improve student achievement and close achievement gaps by focusing resources on students who have yet to attain proficiency or are at risk of falling below proficiency. It will give schools and districts an immediate incentive to identify students who start out far behind and launch them on an accelerated path to proficiency in later grades. It will also compel schools and districts to catch proficient and advanced students whose performance is declining over time. Schools will be encouraged to differentiate instruction and interventions based on individual student needs as well as to provide professional development and technical assistance to educators to assist them with these tasks. Without the use of this growth measure, schools may receive an inaccurate picture of student achievement. The collective use of “status,” “improvement” and “growth” measures yield a comprehensive and complete picture of achievement in today’s schools for community members, districts, the Pennsylvania Department of Education and the United States Department of Education.

Pennsylvania Growth Model Amendment Scheme

Currently Approved	Proposed
STATUS Measures: o AYP Performance Target Current Year o Performance Target 2 Year Average o Performance Target with Confidence Interval o Performance Target 2 Year Average with Confidence Interval	STATUS Measures: o AYP Performance Target Current Year o Performance Target 2 Year Average o Performance Target with Confidence Interval o Performance Target 2 Year Average with Confidence Interval
IMPROVEMENT Measures: o Safe Harbor (Proficiency defined by status) o Reduce the # of students not-proficient from previous year by 10% o Safe Harbor with Confidence Interval o Pennsylvania Performance Index (PPI)	IMPROVEMENT Measures: o Safe Harbor (Proficiency defined by status) o Reduce the # of students not-proficient from previous year by 10% o Safe Harbor with Confidence Interval o Pennsylvania Performance Index (PPI)
	GROWTH Measure: **PROPOSED** o AYP Performance Target Current Year using proficiency by projection to proficiency o AYP Performance Target Current Year using proficiency by projection with a 68% confidence interval

See Appendix E for details on each AYP calculation

Timeline for Implementation of Proposed Amendment

If approved, this growth model amendment would be available to determine the AYP status for all Pennsylvania districts and schools for academic year SY06-07.

Growth Model Integrity

The Pennsylvania Value-Added Assessment System (PVAAS) is a secure, web-based data tool that provides accurate measures of the growth of students, subgroups and schools. PVAAS is based on a high quality statistical analysis of the Pennsylvania System of School Assessment (PSSA) scores for all students in the Commonwealth. This methodology has been independently reviewed and validated by research institutions such as the RAND Corporation (2003) and Tucker and Stronge (2005). The projection metric yielded from PVAAS analysis operates only on student-level achievement data and does NOT bias the analyses with any student demographics or school characteristic data.

The projection feature of PVAAS provides an analytical estimate of each student's trajectory to proficiency. Since the projection targets the goal of proficiency, it focuses on sufficient growth to achieve proficiency. This reinforces and supports USED's understanding that "one year progress for one year of instruction" is not sufficient unless the trajectory to proficiency is confirmed (Spellings, 2005). Pennsylvania's model supports Secretary Spellings' higher standards for achievement as defined in No Child Left Behind.

Conclusion

NCLB requires that "adequate yearly progress shall be defined by the State in a manner that "… results in continuous and substantial progress for all." Pennsylvania believes that a growth measure like PVAAS to measure and document progress toward proficiency is consistent with this requirement and philosophy as a further option for meeting AYP for all schools and districts in Pennsylvania. By measuring the growth of students currently scoring below proficient, especially in mathematics and reading, the efforts of the teachers and administrators who strive to provide access to proficiency are recognized. Acceptance of this proposal not only validates this recognition but also publishes a strong message to educational professionals that growth coupled with status measures is critical for ensuring the success of all of Pennsylvania's students. The Pennsylvania Department of Education believes that its' strong history of commitment, both fiscally and philosophically, to the use of a growth model as a tool for school improvement establishes Pennsylvania as a leading candidate for participation in the USED pilot program of growth models for AYP. The remainder of this proposal specifically addresses USED's Core Principles for a growth model and Pennsylvania's plan to comply with these principles.

Response to USED Seven Core Principles

Core Principle 1: 100% Proficiency by 2014 and Incorporating Decisions about Student Growth into School Accountability "The accountability model must ensure that all students are proficient by 2013-14 and set annual goals to ensure that the achievement gap is closing for all groups of students." (Secretary Spellings' letter, 11/21/05)	
Peer Review Questions	**Pennsylvania Response**
1.1 How does the State accountability model hold schools accountable for universal proficiency by 2013-14? **1.1 Peer Review Probe Questions**	
1.1.1 Does the State use growth alone to hold schools accountable for 100% proficiency by 2013-14? If not, does the State propose a sound method of incorporating its growth model into an overall accountability model that gets students to 100% proficiency by 2013-14? What combination of status, safe harbor, and growth is proposed? Indicate which of the four options listed below is proposed to determine whether a school makes adequate yearly progress (AYP) and for identifying schools that are in need of improvement, and explain how they are combined to determine AYP: 1. Growth alone *2.* Status and growth 3. Status, safe harbor, and growth 4. Safe harbor and growth The Department is planning to evaluate the use of growth models. Once implemented, States participating in the	1.1.1 Pennsylvania proposes to use status, safe harbor/improvement *and* growth to hold subgroups, schools and districts accountable for meeting the goal of 100% proficiency in mathematics and reading by 2013-14. Presently the Pennsylvania Accountability Workbook includes status and safe harbor/improvement measures. The State proposes to use a projection model, rather than a value-added or other form of growth model to evaluate the student academic progress of all students in reading and math towards state standards as an additional alternative for meeting AYP proficiency requirements. Pennsylvania understands that if approved for the growth model AYP pilot, the state will provide data, as directed by USED, showing how the model compares to the current AYP methods.

<table>
<tr>
<td>growth model pilot project will be expected to provide data showing how the model compares to the current AYP status and safe harbor approaches.

➢ What are the grade levels and content areas for which the State proposes to measure growth (e.g., from 2004-05 to 2005-06 in reading and mathematics for grade levels 3-8)?

➢ If the State does not propose to implement its Growth model in all grade levels 3-8 and high school and for both subjects, where are the gaps in Growth Model decisions and what are the implications of those gaps for school accountability?</td>
<td>➢ The State will apply its projection model to both reading and mathematics for students in grades 4-8.

➢ Pennsylvania will not apply its projection model to grades 3 and 11 since in both cases, there will not be two consecutive years of data with which to complete the calculation. For grades 3 and 11, current observed test scores will be used for AYP determination.</td>
</tr>
<tr>
<td>1.2 Has the State proposed technically and educationally sound criteria for “growth targets”[1] for schools and subgroups?

1.2 Peer Review Probe Questions
1.2.1 What are the State’s “growth targets” relative to the goal of 100% of students proficient by 2013-14? Examine carefully what the growth targets are and what the implications are for school accountability and student achievement.

➢ The State should note if its definition of proficiency includes “on track to be proficient” or a related growth</td>
<td>1.2 Under the projection model, a student is considered “proficient” in math or reading if the student is projected to be proficient in the subject in the grades specified in Table 1:

Table 1: Present and Projected Grades
<table><tr><td>Present Grade</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td></tr><tr><td>Projection to Grade</td><td>6</td><td>7</td><td>8</td><td>8</td><td>11</td></tr></table>
The projection model will include all students tested under the Pennsylvania System of School Assessment (PSSA).
➢ A 4th or 5th grade student will be considered proficient if the student is projected to score above the proficiency</td>
</tr>
</table>

[1] “Growth target” denotes the level of performance required in order to meet Apothem State may propose different “growth targets” for reading/language arts and mathematics, different grade spans, etc. This document uses the term “growth target” to try to minimize confusion with “expected growth,” “projected growth,” “growth expectations,” and other terms used in Value-Added and other student longitudinal growth approaches that denote an empirically derived student performance score not necessarily related to the NCLB policy goals of proficiency.

concept. For example, a State may propose that a student who is not proficient in the current grade must be on track to proficiency within three years or by the end of the grade span (e.g., elementary). ➢ A growth model that only expects "one year of progress for one year of instruction" will not suffice, as it would not be rigorous enough to close the achievement gap as the law requires.	standard on the PSSA assessment two years into the future. A 4th or 5th grade student will be considered below proficient if the student is projected to score below the proficiency standard on the PSSA assessment two years in the future. ➢ A 6th or 7th grade student will be considered proficient if the student is projected to score above the proficiency standard on the Grade 8 PSSA assessment. A 6th or 7th grade student will be considered below proficient if the student is projected to score below the proficiency standard on the Grade 8 PSSA assessment. ➢ An 8th grade student will be considered proficient if the student is projected to score above the proficiency standard on the Grade 11 PSSA assessment. An 8th grade student will be considered below proficient if the student is projected to score below the proficiency standard on the Grade 11 PSSA assessment. ➢ Students in their first tested year in Pennsylvania, including 3rd grade students and students with no prior test score, will be considered proficient if they score above the proficiency standard in the current year and considered below proficient if they score below the proficiency standard in the current year. ➢ An 11th grade student will be considered proficient if they score above the proficiency standard in the current year and considered below proficient if they score below the proficiency standard in the current year. ➢ Students who take the Pennsylvania Alternate Student Assessment (PASA) will be considered proficient if they

	score above the proficiency standard in the current year and considered below proficient if they score below the proficiency standard in the current year. These criteria set a short-time horizon for students to attain proficiency. The model expects that each student will make more than one year's worth of progress. By expecting students in greatest need to make the most progress, the proposed model will drive the elimination of student achievement gaps.
1.2.2 Has the State adequately described the rules and procedures for establishing and calculating "growth targets"?	1.2.2 The previously designated Annual Measurable Objectives and Intermediate Goals for mathematics and reading outlined in the Accountability Workbook will continue to set the proficiency targets for all of Pennsylvania's students. These targeted progressions are designed to keep Pennsylvania on track to meet the NCLB 100% proficiency goals for 2013-14.
1.3 Has the State proposed a technically and educationally sound method of making annual judgments about school performance using growth? **1.3 Peer Review Probe Questions**	
1.3.1 1.3.1 Has the State adequately described how annual accountability determinations will incorporate student growth?	1.3.1 Schools and districts meet AYP proficiency requirements if *the district and school and all subgroups* meet the annual measurable objective in *both reading and mathematics* either by meeting the status or improvement measures in the current approved Pennsylvania Accountability Workbook or through the proposed

projection model. The proposed projection model will be applied for either mathematics or reading subject areas or both depending on the outcome of the status and improvement measure AYP determination. When using the proposed projection model to determine AYP:

1. Each district, school or subgroup's projected percentage of students who score proficient or advanced on reading meets the approved annual measurable objective for reading; or

2. Each district, school, or subgroup's projected percentage of students who score proficient or advanced on mathematics meets the approved annual measurable objective for mathematics.

A. Has the State adequately described and provided a rationale for how Annual Measurable Objectives (AMOs) or other criteria for growth would be determined? Has the State provided a table giving the values for the AMOs from the first year the growth model will be applied (e.g., 2005-06) through 2013-14 that includes rigorous increases in school performance throughout that time? Does the model set reasonable, challenging, and continuously improving annual expectations for student growth?
   - "Growth models that rely on substantial increases in the growth rates of students or schools in the last few years are not acceptable, but the Department is

A. The table below specifies those AMO targets. These targets are not changed because of this proposed amendment.

Year	2002-2004	2005-2007	2008-2010	2011	2012	2013	2014
Percent Proficient in Reading	45	54	63	72	81	91	100
Percent Proficient in Math	35	45	56	67	78	89	100

open to models that set a point in time as the goal (e.g., end of grade in a particular school; within four years).In setting these standards, the State should demonstrate how accountability is distributed among all the grades and not postponed to this point in time. The Department is concerned that if the State's Growth Model allows attainment of the proficiency standard by individual students to be delayed or is tied to standards that are not considerably more rigorous with each consecutive grade, then it becomes too easy to minimize or delay the importance of accelerated growth" (Secretary Spellings' letter, 11/21/05).	
B. For any proposed confidence intervals or other statistical methods to be applied to the decision about meeting the AMO for growth, has the State clearly described the rationale for the use of the specific statistical method (including minimum group size and any multi-year averaging), and the procedures for applying the method?	B. Pennsylvania's accountability system values making accurate decisions about schools. An accountability system that makes accurate, fair and equitable decisions reflects the values of Pennsylvania educators, parents, business partners and community persons by relating accountability to teaching and learning. This value does not change with the implementation of a growth model. Statistical safeguards, such as minimum N, and confidence intervals are used to promote consistency, fairness, and equity in accountability decisions for schools by balancing reliable and valid decision making with inclusion. The projection model will use all current rules approved under

	Pennsylvania's Accountability Workbook, including disaggregating by subgroup, counting only students with full academic year status, and applying a minimum subgroup size of 40. In order to assure the statistical validity and reliability of AYP decisions based on the projection model, Pennsylvania requests the use of a 68% confidence interval around the percent of students considered proficient by projection to proficiency. This request is designed to protect against Type I error by recognizing the sampling error inherent in such a system. The intent of any accountability system under NCLB, whether it is status or growth, is to use information gathered from one set of students tested to make determinations about the school as a whole. Generalizations from one cohort to a larger population are subject to sampling error. In addition, because the growth model scores are based on score projections, which have errors marginally higher than observed scores, the use of confidence intervals to correct for measurement error is even more critical. To ensure that accurate judgments continue to be made about schools, Pennsylvania requests the use of a 68% confidence interval around the percent of students considered proficient by projection to proficiency. Pennsylvania believes that this request is consistent with the use of confidence intervals for the methods of meeting AYP by status measures.
C. For future evaluation purposes, does the State's proposal provide evidence of the validity and reliability of the proposed growth model, including impact of	C. Pennsylvania uses a longitudinal data structure with accepted, reviewed and recognized statistical methodology to calculate its projections (mixed effect multi-variate longitudinal modeling).

use/non-use of the growth model on validity and reliability of overall school accountability judgments?	This methodology has been validated and widely used throughout the country for more than a decade. Independent review of the methodology is included in the references in Appendix C.
1.3.2 Has the State adequately described how it will create a unified AYP judgment considering growth and other measures of school performance at the subgroup, school, district, and state level?	1.3.2
A. Has the State proposed a sound method for how the overall AYP judgment (met/not met) for the school will be made, incorporating judgment of student growth?	A. The AYP Calculation Progression Table on page eight of the proposal outlines the proposed methods of meeting AYP. A school or district will make AYP if it meets all proficiency requirements of the status, safe harbor/improvement or projection model, meets the 95% participation rate for all subgroups, and meets the additional indicator.
B. Has the State proposed a sound method for how the overall AYP judgment for the school will incorporate growth in subgroup performance? ➢ Are the method and criteria for determining subgroup performance on growth the same as for students in the school as a whole?	B. A subgroup will make AYP if it meets the proficiency requirements of the status, safe harbor/improvement or projection model and meets the 95% participation rate.
C. Has the State proposed categories for understanding student achievement at the school level and reports for growth performance and AYP judgments that are clear and understandable to the public?	C. Pennsylvania will report the results of the status, safe harbor/improvement, and the projection model for all schools and districts in a manner that is clear and understandable to the public. These results will be reported in table and graphic forms to the public via its PAAYP.com website. Additional material

	detailing the projection model will be added to the PAAYP.com website to provide parents with the opportunity to use the information to inform their educational decisions.
1.4 Does the State proposed growth model include a relationship between consequences and rate of student growth consistent with Section 1116 of ESEA? **1.4 Peer Review Probe Questions** 1.4.1 Has the State clearly described consequences the State/LEA will apply to schools? Do the consequences meaningfully reflect the results of student growth? ➢ The proposed interventions must comply with the Section 1116 requirements for public school choice, supplemental educational services, and so on. ➢ If proposed, the State should explain how it plans to focus its school intervention efforts by incorporating the results from a growth model. For instance, a State should be prepared to explain how a school that does not meet either traditional AYP goals or growth-based accountability goals might be subject to more rigorous intervention efforts than schools not making AYP on only one accountability measure.	1.4 1.4.1 Pennsylvania's current Accountability Workbook designates a school improvement process for those schools not meeting AYP targets. The process complies with the Sec. 1116 requirements. Under the proposed amendment, the present process for school improvement and consequences for districts and schools would continue. Intervention efforts would proceed as outlined in the current Pennsylvania Accountability Workbook. ➢ The projection to proficiency model will allow Pennsylvania to focus interventions and support for districts and schools that are not on accelerated paths to proficiency. By reporting the results of the projection model for all districts, schools and subgroups in elementary/middle schools, the state will publicly recognize schools and district that are successfully placing individual students on accelerated paths to proficiency and catching students at-risk of falling out of the proficient range in the future.

Core Principle 2: Establishing Appropriate Growth Targets at the Student Level "The accountability model must establish high expectations for low-achieving students, while not setting expectations for annual achievement based upon student demographic characteristics or school characteristics." (Secretary Spellings' letter, 11/21/05)	
Peer Review Questions	**Pennsylvania Response**
2.1 Has the State proposed a technically and educationally sound method of depicting annual student growth in relation to growth targets? **2.1 Peer Review Probe Questions** 2.1.1 Has the State adequately described a sound method of determining student growth over time? A. Is the State's proposed method of measuring student growth valid and reliable? 1. Are the "pre-" and "post-" test scores appropriately defined and adequately measured? 2. If the State will not use a single score for pre- and/or post- test scores (e.g., using an aggregation of multiple scores from multiple years), does the State adequately explain and justify how the scores would be combined, what the weights are for each score, and how and whether the scores are/are not comparable across students and across time? 3. Information about the availability and technical quality of proposed data will be considered in Core Principle 5. The probes associated with Principle 2 are focused on how the change in achievement is measured and valued.	2.1.1 A. Pennsylvania's proposed projection model relies on a robust methodology that uses all of a student's prior achievement scores to estimate the student's achievement level at a future point in time. The model estimates a student's performance on a future assessment based on the student's test performance history and the histories of students with similar performance patterns. All available achievement data on the student are used to yield the projection. The only predictor variables are the student's prior test scores. Demographic data are <u>not</u> used in the analyses. The model assumes that the student will have the average Pennsylvania schooling experience in the future. The model then includes estimated mean scores for the average school in the state and regression coefficients is pooled within schools across the state. These coefficients are updated each year to reflect new student cohort test scores at the projection endpoints. This proven model, by using all available achievement data, provides more robust

	projections than simple pretest-post test models. For example, to achieve a 6th grade student's projected score to his/her 8th grade score on the reading test, the statistical methodology uses scores from students who took the reading exam in the current year who have the same historical pattern of test scores as the 6th grade student. If the student has 3rd, 5th, and 6th grade scores (but no 4th grade scores) the methodology estimates regression coefficients for these scores based on the subset of students who took the reading test in the current year who also had 3rd, 5th, and 6th grade scores, but no 4th grade scores. These coefficients are then applied to the individual student's 3rd, 5th, and 6th grade scores to calculate the student's projected score on the 8th grade reading test. If the student has made progress between 3rd and 6th grade, the model will show if this progress has been sufficient to predict that the student will reach proficiency by the time he or she takes the 8th grade exam.
B. Has the State established sound criteria for growth targets at the student level, and provided an adequate rationale? If the State is assigning a value determination at the student level annually with regard to each student's growth, has it used a sound process and assigned specific values for those growth targets? For example, if a State has four performance categories, would movement between each category be weighted equally or would some categories be weighted more heavily than others?	B. The nature of the Pennsylvania amendment precludes the necessity of growth targets since it uses the projection capability of the longitudinal data structure of student achievement data. The longitudinal data structure includes only student achievement data and does not include any data regarding student demographics or school characteristics. ➢ Each student is assigned the designation of proficient or non-proficient based on his/her individual projection to proficiency, independent of the projections of all other students. There is no aggregation of these projections. This analysis is NOT biased by student demographics or

If the State would only calculate "difference" or "change" scores for each student, and then aggregating to the subgroup and/or school levels, then the State should clearly give its rationale in this section. Would the model ensure that student growth expectations are not set or moderated based on student demographics or school characteristics? The model must have the same proficiency expectations for all students, while setting individual growth expectations for students to enable them to meet grade level standards. • If the State proposes a regression or multivariate/multi-level model, the independent variables may not include race/ethnicity, socio-economic status, school AYP status, or any other non-academic covariate. • Does the model establish growth targets in relation to achievement standards and not in relation to "typical" growth patterns or previous improvement, unless there is evidence and a clear rationale that those factors are related to the overall goal of achieving proficiency for all students? • Would gains of high performing students compensate for lack of growth among other students?	school characteristics. Since the projections are determined by a student's past scores compared to students who have the same pattern of test scores, students with similar test performances will have the same projected score. ➢ Students' testing histories will be updated each year as additional test scores become available. This new information will be considered as projections are re-determined and updated. ➢ The proposed projection to proficiency metric does not moderate its expectations for students based on student demographics or school characteristics. Projections are determined by comparing each individual student's testing history to students with similar testing histories, not those with like demographics. Using information from all students with like histories helps to provide more robust projections for each student. • Growth targets are aligned with those achievement standards established in Pennsylvania's Accountability Workbook. • Since the projection is calculated for each student, the effect of either a high performing or low performing student will not be an issue. The projection to proficiency for each student is the calculation for this amendment.
4. Does the State have a plan for periodically	4 Pennsylvania plans to validate both the model and growth targets on an ongoing basis using simulations

evaluating the appropriateness of the student-level growth targets criteria?	and analyses performed by both the statistical vendor that Pennsylvania currently uses to calculate the growth projections and an independent measurement consulting firm.

Core Principle 3: Accountability for Reading/Language Arts and Mathematics Separately "The accountability model must produce separate accountability decisions about student achievement in reading/language arts and in mathematics." (Secretary Spellings' letter, 11/21/05)	
Peer Review Questions	**Pennsylvania Response**
3.1 Has the State proposed a technically and educationally sound method of holding schools accountable for student growth separately in reading/language arts and mathematics? **3.1 Peer Reviewer Probe Questions** 3.1.1 Are there any considerations in addition to the evidence presented for Core Principle 1? ➢ The growth model proposal must include separate decisions for reading/language arts and mathematics, and maintain validity and reliability, minimize measurement error, and support empirical integrity in the accountability system. How does the model achieve these specifications, especially in small schools or schools with high mobility?	3.1 3.1.1 ➢ Under the projection model, Pennsylvania will calculate projected scores separately for reading and mathematics and will base these projections on data from the same content area only. For example, to determine whether a school/district/subgroup met the annual proficiency target in reading/language arts, it will use student projected scores in reading. Pennsylvania's projection methodology is very flexible. It does not require vertically-equated test score scales nor does it assume a specific growth function (see Technical Appendix F

• Does the model include assessments for other content areas (e.g., covariance matrices to estimate student performance or projected performance in a content area)? If so, the State should demonstrate that achievement on those other assessments does not compensate for failure to achieve proficiency in reading/language arts or mathematics.	for the model). In order to increase reliability and dampen measurement error, the projection methodology uses all of a student's prior achievement scores from all PSSA reading and math assessments to project future scores. In small schools and schools with high mobility, projected scores are more valid measures of school performance than current-year scores because they incorporate all of a student's prior achievement data. Pennsylvania's longitudinal database follows students across time and across the Commonwealth, maximizing the reliability of the projections for these schools.

Core Principle 4: Inclusion of All Students

"The accountability model must ensure that all students in the tested grades are included in the assessment and accountability system. Schools and districts must be held accountable for the performance of student subgroups. The accountability model, applied statewide, must include all schools and districts." (Secretary Spellings' letter, 11/21/05)

Peer Review Questions	**Pennsylvania Response**
4.1 Does the State's growth model proposal address the inclusion of all students, subgroups and schools appropriately? **4.1 Peer Review Probe Questions**	
4.1.1 Does the State's growth model address the inclusion of all students appropriately?	4.1.1 All students who qualify for AYP proficiency determinations based on the NCLB criteria for full academic year, as

	detailed in the Pennsylvania Accountability Workbook, will be included in the projection model. If a student does not have a previous PSSA score, his/her current score will be used instead of a projected score. The current-year scores of 3rd grade students and students new to the State will be used in the projection model. The state will include current year scores of students assessed under the PASA. The State will include current-year scores of Grade 11 students.
A. Ideally, every student will have a pre- and a post-score, and a school will be clearly accountable for all students' achievement even when applying the "full academic year" parameters. However, there will be situations in which this is not the case. Are the State's proposed rules for determining how to include student achievement results (when data are missing) in the growth model technically and educationally sound? ➢ For example, if a State proposes to "impute" missing data, it should provide a rationale and evidence that its proposed imputation procedures are valid. A State proposing such a growth model must address how many students would be excluded from its calculations of growth because they lack a score, and provide an acceptable explanation of how these exclusions would not yield invalid or misleading judgments about school performance.	A. Pennsylvania does not impute missing data in the projection model. The model uses all available data on student test performance history to project future performance even when a student does not have a test score in every subject in every year. The PASecureID system (Pennsylvania's state-wide student identifier) will allow Pennsylvania to track students' information across time, across districts and across the state. This will allow Pennsylvania to maximize the number of student projections used in the projection model.

➢ Does the State have an appropriate proposal for including students who participate with alternate assessments and/or alternate/modified achievement standards (in one or more years for calculating growth)? ➢ Does the State's definition of FAY include students appropriately when applied in the growth model context? For example, a State that defines FAY as "participating in the assessment in the same school the previous year" will need to modify that definition for its growth proposal to include students who cross school boundaries over time. ➢ What does the State propose to do to measure academic growth for students in grade three or the initial grade tested? ➢ How does the State propose to distinguish between growth for a student who moves from one grade level to another and growth for a student who is retained in a grade level for two years or is promoted at mid-year? B. What other strategies will the State use to include, in its NCLB accountability system, students who might be excluded from the growth model calculations?	B. If a student does not have a previous PSSA score, his/her current score will be used instead of a projected score.
4.1.2 Does the State's growth model address the inclusion of all subgroups appropriately? A. States must ensure that student subgroups are neither systematically nor inadvertently excluded from	4.1.2 A. The projection model holds schools accountable for the achievement of all subgroups in both reading and math. All subgroups must meet the AMO in the content area for that year.

participation in the growth model; the model cannot eliminate or minimize the contribution of each subgroup. Are the State's proposed rules for determining how to include subgroup accountability in the growth model technically and educationally sound? ➢ Has the State adequately addressed implications of its proposed growth model for subgroup inclusion in addition to that in Core Principle 1? (For example, has it addressed "minimum group-size" requirements for subgroups?) ➢ Does the State have an appropriate proposal for including students who change subgroup classification over the time period when growth is calculated (e.g., LEP to non-LEP)? ➢ If applicable, how does the State proposal address the needs of students displaced by Hurricanes Katrina and Rita? For example, how does the proposal interact with State plans, if any, to develop a separate subgroup of displaced students, consistent with the Secretary's guidance of Sept. 29, 2005?	➢ Since AYP determination occurs each year for students as identified on the current state assessment, a change in subgroup status will not affect the calculation of individual student projections. Student scores, whether current or projected, will be assigned to the subgroup to which the student belongs in the current year. ➢ There is no subgroup in Pennsylvania to address the needs of students displaced by Katrina or Rita.
4.1.3 Does the State's growth model address the inclusion of all schools appropriately? A. Does the State provide an adequate plan and rationale for how the system will be applied to all schools consistently across the State to yield an AYP determination each year? Has the State adequately	4.1.3 A. All Pennsylvania schools and districts receive an AYP determination each year. The projection model will be included as an additional AYP measure applied to all students in these determinations.

described and provided a rationale for any proposed exceptions? ➢ The State may propose to apply the growth model only to schools with adequate assessment data. If that is the case, it should propose how other schools, such as K-2 schools, single-grade schools, and high schools, will be held accountable (e.g., through continuing its approved statutory AYP/safe harbor accountability system for those schools). ➢ The State should propose how it will deal with common conditions that would preclude the calculation of a growth score (e.g., school boundary changes, school closings, new schools, grade reconfiguration). ➢ How would the model ensure that all schools are accountable for student achievement, even when the number of tested students in the school is small or constantly changing?	➢ Beginning in 2006-07, the proposed amendment will be implemented for all districts in Pennsylvania. Pennsylvania's uniform statewide identifier (PASecureID) for students addresses the issues of mobility and changes in district/school conditions. The data for each student will be archived in a unique test history record that is independent of school attendance or school changes. ➢ The projection model will adhere to the existing provisions outlined in the Accountability Workbook; specifically AYP proficiency determinations will not be calculated for subgroups with fewer than 40 students. In addition, the statistical methodology for projecting performance on a future assessment requires a minimum of 10 students per grade level. Projections will not be calculated for schools that do not meet this criterion. In these very rare cases or for schools that do not contain any of the PSSA tested grades, accountability determinations will be determined as specified in the Accountability Workbook.

Core Principle 5: State Assessment System and Methodology

"The State's NCLB assessment system, the basis for the accountability model, must include annual assessments in each of grades three through eight and high school in both reading/language arts and mathematics, must have been operational for more than one year, and must receive approval through the NCLB peer review process for the 2005-06 school year. The assessment system must also produce comparable results from grade to grade and year to year." (Secretary Spellings' letter, 11/21/05)

Peer Review Questions	Pennsylvania Response
5.1 Has the State designed and implemented a Statewide assessment system that measures all students annually in grades 3-8 and one high school grade in reading/language arts and mathematics in accordance with NCLB requirements for 2005-06, and have the annual assessments been in place since the 2004-05 school year? **5.1 Peer Review Probe Questions <ED STAFF TO COMPLETE>**	5.1
5.1.1 Provide a summary description of the Statewide assessment system with regard to the above criteria. ➢ For both 2004-05 and 2005-06, did the State implement an assessment system that measures State adopted content standards in reading/language arts and mathematics? ➢ Did the State produce individual student, school, and district test results for both years?	5.1.1 ➢ Pennsylvania fully implemented its state system of assessment in grades 3-8 and 11, called the Pennsylvania System of School Assessment (PSSA) in 2005-2006. Mathematics and reading have been assessed in grades 5, 8 and 11 for many years and the State was compliant with the assessment requirements of IASA. Mathematics and reading assessments in grade 3 have been in place for four years. All of the state assessments are aligned with the Pennsylvania State Content Standards for reading and math. The spring of 2007 will mark the 2nd administration of PSSA in all grades. The State produces student, school and district reports for all state assessments. The alignment and rigor of the PSSA was independently verified by HUMRO in 2004 and Achieve in 2004 and 2005.
5.1.2 Has the State submitted its statewide assessment system for NCLB Peer Review and, if so, was it approved for 2005-06?	5.1.2 Pennsylvania submitted its statewide assessment system for NCLB Peer Review in May 2006 and received the

➢ If it was not fully approved, what are the deficiencies and to what extent will they affect the State's ability to measure growth in each subject? ➢ If the State has not yet received approval of its assessment system, when does the State plan to submit evidence of compliance with the NCLB standards and assessment requirements?	classification "Approval Pending – 2 Issues". However, PDE has addressed all peer review findings as outlined in Dr. Johnson's letter and will be submitting this evidence by January 15, 2007 after incorporating the comments and suggestions from the State's technical advisory committee at the conclusion of the November 30-December 1, 2006 meeting. A detailed plan and timeline has been submitted (on 7.31.06) and has been approved by USED. This plan addresses all outstanding issues to receive a Full Approval from USED following a January 2007 Peer Review. The issues addressed in the detailed plan and timeline will not impact our ability to measure growth. The "outstanding" issues include a comparability study for the Spanish math test, a monitoring system, and technical documentation regarding the standard setting, as well as evidence of the Board adoption of the standards and cut scores.
5.2 How will the State report individual student growth to parents? **5.2 Peer Review Probe Questions** 5.2.1 How will an individual student's academic status be reported to his or her parents in any given year? What information will be provided about academic growth to parents? Will the student's status compared to the State's academic achievement standards also be reported?	5.2 5.2.1 Each school district in Pennsylvania receives information on student projections linked to future assessments via a secure website. This information compares each student's projected performance to the state's standards for proficient on future state assessments. Access to the website occurs through a secure login and password made available to the superintendent of each school district. The superintendent may then assign access to additional users. A printable version of each student's projection information is available

	for schools to share with parents along with their student's status measures. Pennsylvania will continue to support local districts in their communication with their constituencies through a variety of resources and professional development opportunities. Other methods currently in place to provide parents with information about an individual student's, school's, or district's status in relationship to Pennsylvania standards and AYP targets will continue to be provided to the parents via methods such as www.paayp.com and the Grow Reports.
5.3 Does the Statewide assessment system produce comparable information on each student as he/she moves from one grade level to the next? **5.3 Peer Review Probe Questions** The State assessment system – that is the achievement levels and content expectations – needs to make sense from one grade to the next, and even within achievement levels for it to support a growth model. These probes will help the peers understand the assessment system's capability for use in growth models.	5.3 In 2006, Pennsylvania instituted PASecureID, a unique student identifier that enables the Department of Education to produce comparable information on each individual student. A variety of methods are utilized to track and compare student levels from year to year, including but not limited to the eMetric Data Interaction website. Pennsylvania has established a carefully articulated system of content expectations and achievement standards to support content-based inferences of student progress across grade levels. In June, 2006, the State Board adopted the Performance Level Descriptors and the PSSA Performance Level Cut Scores for grades 4, 6, and 7 in mathematics and reading to be used with previously established Performance Level Descriptors and Cut Scores for grades 3, 5, 8 and 11. An up-to-date technical guide on the state assessment is

	published each year and is available for further information.
5.3.1 Does the State provide evidence that the achievement score scales have been equated appropriately to represent growth accurately between grades 3-8 and high school? If appropriate, how does the State adjust scaling to compensate for any grades that might be omitted in the testing sequence (e.g., grade 9)? Did the State provide technical and statistical information to document the procedures and results? Is this information current?	5.3.1 PSSA does not use a vertically-equated score scale across grades. There are many legitimate criticisms of vertical score scales (e.g., Hill, et al., Lissitz & Huynh, 2003) and PDE felt that a system of vertically-articulated achievement standards would lead to more valid inferences about growth than vertical score scales. Pennsylvania publishes and makes available on the Department of Education website, technical statistical information on the PSSA for each year of its administration. Technical documents are produced immediately after release of scores. Technical analyses are available at http://www.pde.state.pa.us/
5.3.2 If the State uses a variety of end-of-course tests to count as the high school level NCLB test, how would the State ensure that comparable results are obtained across tests?[Note: This question is only relevant for States proposing a growth model for high schools and that use different end-of-course tests for AYP.]	5.3.2 Pennsylvania does NOT use end of course tests for AYP determinations.
5.3.3 How has the State determined that the cut-scores that define the various achievement levels have been aligned across the grade levels? What procedures were used and what were the results?	5.3.3 PDE and its test contractor—Data Recognition Corporation—with advice from the technical advisory committee designed the standards validation procedure to maximize the likelihood that the content and normative interpretations will lead to valid inferences about student progress across grades

	relative to the underlying content demands. An extensive set of methods, based on a modified Bookmark procedure, were used to ensure that the meaning of the various cut scores across grades, particularly the proficient score, were well articulated in terms of the content standards. Among other strategies, the technical consultants suggested "starting points" to focus the Bookmark deliberations and relied on cross-grade level discussions to help articulate the cut scores across grades.
5.3.4 Has the State used any "smoothing techniques" to make the achievement levels comparable and, if so, what were the procedures?	5.3.4 Finally, statistical smoothing approaches—approved by the technical advisory committee—were used to ensure the comparability of the achievement levels. Details are available in the technical analyses available on the Department of Education website at http://www.pde.state.pa.us/

5.4 Is the Statewide assessment system stable in its design? **5.4 Peer Review Probe Questions**	5.4
5.4.1 To what extent has the statewide assessment system been stable in its overall design during at least the 2004-05 and 2005-06 academic terms with regard to grades assessed, content assessed, assessment instruments, and scoring procedures?	5.4.1 Statewide assessments in mathematics and reading for grades 3, 5, 8 and 11 have been stable for several years. Assessments in grades 4, 6, and 7 were field tested in 2004-05 and are now fully operational. No changes are anticipated for the testing system in the foreseeable future.
5.4.2 What changes in the statewide assessment system's overall design does the State anticipate for the next two academic years with regard to grades assessed, content assessed, assessment instruments, scoring procedures, and achievement level cut-scores? ➢ What impact will these changes have on the State's proposed growth model? How does the State plan to address the assessment design changes and maintain the consistency of the proposed growth model?	5.4.2 Pennsylvania will continue with its current statewide assessment system regarding mathematics and reading for the foreseeable future. No changes are anticipated to Pennsylvania's assessment system that would require adjustments to be made in the proposed growth model.

Core Principle 6: Tracking Student Progress "The accountability model and related State data system must track student progress." (Secretary Spellings' letter, 11/21/05)	
Peer Review Questions	**Pennsylvania Response**
6.1 Has the State designed and implemented a technically and educationally sound system for accurately matching student data from one year to the next?	6.1

6.1 Peer Review Probe Questions	
6.1.1 Does the State utilize a student identification number system or does it use an alternative method for matching student assessment information across two or more years? If a numeric system is not used, what is the process for matching students?	6.1.1 Pennsylvania received a three-year, $4 million USED grant to implement a longitudinal student data system. This includes the implementation of a unique statewide student identifier-PASecureID. Pennsylvania implemented a growth model in 2002 that uses district-level identifiers. The PASecureID ensures the integrity of longitudinal merging needed to yield a data metric for AYP. This was implemented in SY05-06.
6.1.2 Is the system proposed by the State capable of keeping track of students as they move between schools or school districts over time? What evidence will the State provide to ensure that match rates are sufficiently high and also not significantly different by subgroup?	6.1.2 The PASecureID system will enable accurate tracking of students across districts. Match rates for grades 4 and 6 from the 2006 test administration are contained in Appendix G.
6.1.3 What quality assurance procedures are used to maintain accuracy of the student matching system?	6.1.3 The PASecureID system will include quality assurance checks and procedures.
6.1.4 What studies have been conducted to demonstrate the percentage of students who can be “matched” between two academic years? Three years or more years?	6.1.4 Pennsylvania’s experience with using district-level identifiers demonstrates its capacity to yield growth measures and track students over time. SAS, Inc does a multi-field merge of historical student data to yield reporting (see Appendix G).
6.1.5 Does the State student data system include information indicating demographic characteristics (e.g., ethnic/race	6.1.5 The Pennsylvania longitudinal data structure will contain data

category), disability status, and socio-economic status (e.g., participation in free/reduced price lunch)?	regarding all NCLB student demographic characteristics. These data are used only for the purpose of merging student records and reporting results for AYP.
6.1.6 How does the proposed State growth accountability model adjust for student data that are missing because of the inability to match a student across time or because a student moves out of a school, district, or the State before completing the testing sequence?	6.1.6 The projection methodology used in Pennsylvania is remarkable in that it utilizes student records with similar testing performance histories in its modeling process. Therefore, the issue of "fractured records" or missing data is NOT of any consequence. If a student does not have the minimum number of PSSA data points (three), the projection model will use the student's current-year score. For a further discussion of the projection methodology, see the Technical Appendix E.
6.2 Does the State data infrastructure have the capacity to implement the proposed growth model? **6.2 Peer Review Probe Questions**	6.2 The capacity for storage, merging and analyses for the proposed growth model exceeds the requirements for this proposal now and for the future. A statewide longitudinal data structure was established for PVAAS in 2002.
6.2.1 What is the State's capability with regard to a data warehouse system for entering, storing, retrieving, and analyzing the large number of records that will be accumulated over time?	6.2.1 Pennsylvania had demonstrated its capability of managing longitudinal data structures. For the past two years, the Pennsylvania Department of Education has been using other longitudinal data systems to report multiple years of individual student learning with the PSSA Data Interaction System and the Pennsylvania Academic Reports by GROW. For further information please see http://www.paayp.com.
6.2.2 What experience does the State have in analyzing longitudinal data on student performance?	6.2.2 Pennsylvania has been involved with longitudinal data structures and their analyses for growth since 2002 with 100 pilot districts and since SY05-06 with all 501 school districts.

	These analyses included value-added reports for school improvement and student projections of performance on future assessments, the proposed metric for this proposal.
6.2.3 How does the proposed growth model take into account or otherwise adjust for decreasing student match rates over three or more years? How will this affect the school accountability criteria?	6.2.3 Pennsylvania believes that the successful implementation of PASecureID will adequately address the issue of decreasing match rates. Pennsylvania's four-year experience with the 100 pilot districts and recent expansion to all 501 school districts verifies that a unique identifier will be effective in this regard. This will be reassessed annually. Should decreasing match rates disable the option of calculating projections for a particular student, the student's status measures will be used for AYP purposes.

Core Principle 7: Participation Rates and Additional Academic Indicator The accountability model must include student participation rates in the State's assessment system and student achievement on an additional academic indicator. (Secretary Spellings' letter, 11/21/05)	
Peer Review Questions	**Pennsylvania Response**
7.1 Has the State designed and implemented a Statewide accountability system that incorporates the rate of participation as one of the criteria?	7.1 The projection model only applies to reading and math proficiency. Schools and districts with subgroups that do not meet the 95% participation rate or the other indicator requirements will not make AYP.
7.1 Peer Review Probe Questions	
7.1.1 How do the participation rates enter into and affect the growth model proposed by the State?	7.1.1 Pennsylvania currently has "participation" as one variable in their accountability plan and will continue to include this data as a measure of accountability as defined in the current

	approved Pennsylvania Accountability Workbook.
7.1.2 Does the calculation of a State's participation rate change as a result of the implementation of a growth model?	7.1.2 The proposed amendment does not change the participation policy accepted in the current approved Accountability Workbook by USED.
7.2 Does the proposed State growth accountability model incorporate the additional academic indicator? **7.2 Peer Review Probe Questions**	
7.2.1 What are the "additional academic indicators" used by the State in its accountability model? What are the specific data elements that will be used and for which grade levels will they apply?	7.2 All academic indicators included in Pennsylvania's approved Accountability Workbook will continue to be followed. Pennsylvania's proposed amendment incorporates no additional academic indicators.
7.2.2 How are the data from the additional academic indicators incorporated into accountability determinations under the proposed growth model?	

Appendix A - History of Growth Models in Pennsylvania

While the use of a growth model as a part of a state's accountability plan for NCLB has only recently become open for national discussion, Pennsylvania has been implementing the use of a growth metric at a state level for the past four years. The Pennsylvania Department of Education has taken a national lead in the use of a growth model as part of its ongoing statewide school improvement process. The Pennsylvania Value-Added Assessment System (PVAAS), the Pennsylvania Department of Education's proposed growth model, is a system that began pilot implementation in 2002 and has evolved to statewide implementation during the 2005-2006 school year. One of the unique features of PVAAS is the capability of calculating projections to proficiency for students in the longitudinal database without bias for student demographics or school characteristics, unlike other value-added models.

Initially recommended to PDE by the Pennsylvania League of Urban Schools (PLUS) in 2002, it was reviewed by a No Child Left Behind statewide committee of practitioners who recommended it as part of the Pennsylvania Accountability Workbook for NCLB. The Pennsylvania State Board of Education subsequently passed a resolution to adopt "a value-added approach across the Commonwealth." In 2002, the PVAAS model was selected because of the rigorous statistical methodology used to yield value-added analysis. The PVAAS model uses all available student data as part of its analysis and does not allow schools or districts to "make excuses" about achievement due to demographics and/or school characteristics. The excuses for low achievement are removed with the PVAAS model of growth analyses.

Initially, PVAAS operated as a pilot with 100 participating school districts, including the Pittsburgh Public School District and the School District of Philadelphia. These districts supplied archived data from standardized testing from past years to yield growth reports. As Pennsylvania moved into its statewide implementation, school districts received a value-added analysis using the PSSA performances of their students in grades 3 through 8 and projections to performance in future grades. Beginning in the fall of 2006, Pennsylvania districts received one of two types of value-added assessment reports: the 100 Pilot school districts received full reporting for grades 4 through 8 and 11; the other 401 school districts received value-added reports for grades 4 and 6. In the fall of 2007, all 501 districts will receive value-added reports at the district, school, subgroup and student level for grades 4 through 8 and 11. These reports are web-based via a secure password for the superintendent and easily accessible to the designated school personnel to use as part of a continuous school improvement process. In addition, beginning in the fall of 2006, projections of future performance using the PVAAS data were available to all districts and schools.

Pennsylvania has consistently sought to include a wide range of stakeholders as it developed its plan for the use of PVAAS statewide. In October 2004, the Pennsylvania Secretary of Education, Gerald L. Zahorchak D.Ed., formed a statewide work group of over 50 education, legislative, association, parent advocacy, and business leaders to assist with the implementation of PVAAS in Pennsylvania (See Appendix D for list of workgroup

members). This group provides ongoing advisement to the Department regarding communications about PVAAS, as well as statewide implementation and professional development. This group provided support and feedback in the development of Pennsylvania's proposal to USED for a growth model as a part of its Accountability Workbook. Through the use of this powerful partnership, the Pennsylvania Department of Education has sought to strengthen the use of PVAAS in accomplishing their shared goals of improved student achievement and progress.

Successful implementation of a growth measure is much more than the provision of a growth metric. Communications, professional development and technical assistance are essential components for successful implementation. Pennsylvania has three years of experience of providing statewide communications, extensive statewide professional development and district/school-level technical assistance to its schools as part of the PVAAS pilot program. Pennsylvania had resources allocated and committed to support these areas for statewide implementation for SY02-03, SY03-04, SY04-05, as well as SY05-06 and SY06-07.

- Communication Materials: The Pennsylvania Department of Education has produced a series of communication materials designed to provide an overview of PVAAS. This communications kit is available on the PDE website to stakeholders across the Commonwealth. These materials have been disseminated to all 501 district superintendents, curriculum directors, special education coordinators, over 3,000 building principals, and intermediate units, as well as statewide organizations. The communication kit includes the following materials: *Introductory Guide for Pennsylvania Educators, PVAAS History, Intent and Timeline for Implementation Tri-fold, PVAAS Pilot Districts/Map, Educator Testimonials, Value-Added FAQ, PVAAS FAQ, Press Release for Districts, Tips on Communicating with Your School Board and Community about PVAAS, Benefits for District Administrators, Benefits for Building Principals, Benefits for Teachers and Pilot District Stories*
- Professional Development: A cadre of materials that assist with navigation of the PVAAS website, interpretation of PVAAS reports and the integration of PVAAS with other statewide initiatives were provided to all districts/school in Pennsylvania.
- Technical Assistance: A statewide core team of consultants was extensively trained in the methodology of PVAAS and the interpretation of PVAAS reports. This team has provided district/school level technical assistance in the interpretation and use of the PVAAS growth measures. In SY05-06 this core team provided "train the trainer" workshops to intermediate units and school districts in the principles and use of the PVAAS model.
- Communications: In addition to supporting those participating in the PVAAS pilot, the core team has presented extensively at national, state and local conferences sponsored by professional organizations on the principles of PVAAS.

This extensive commitment to the understanding and use of a growth model allows Pennsylvania a unique "head start" in formalizing the concept of growth paired with its accountability plan. Pennsylvania is poised and willing to serve as a national model for the comprehensive use of a growth model as a part of NCLB.

Appendix B - Integration of Growth Model

Pennsylvania has actively been pursuing the integration of assessment data for instructional decision-making since 2001. Through the Pennsylvania Department of Education's efforts, the following data tools continue to be implemented throughout the Commonwealth:

- PVAAS – The Pennsylvania Value-Added Assessment System provides insight into the academic growth of students in addition to the status measures provided by the PSSA and provides projections to proficiency of future performance for all students.
- Assessing to Learn: Pennsylvania Benchmark Assessment System – 4Sight Benchmark assessments focusing on achievement of proficiency on the Pennsylvania State Standards in mathematics and reading have been implemented in over 300 districts (approximately 18,000 classrooms) in grades 3 through 11.
- eMetric – A sophisticated, secure web-based data tool for analyzing status measures from the PSSA has been available to districts.
- PA AYP – This web-based tool allows access to AYP data for all educators and the general public for all districts and schools in the Commonwealth.
- Pennsylvania Achievement Reports –A secure web-based data tool available to educators containing status data with recommendations and support materials for school use. A non-secure version can be accessed by the general public and includes information for parents and the community.

Pennsylvania is striving to support districts in using an integrated implementation of all of these sources of status, improvement and growth data to inform curricular and instructional decision-making focused on the achievement of proficiency of all students. Pennsylvania has identified that PVAAS is an established and reliable measure of academic growth that provides an important complement to status measures in this regard. The use of both status and growth measures has demonstrated benefits in deliberations for educational improvements for all students at all levels and stages of development. Pennsylvania is integrating PVAAS into the following statewide improvement efforts:

- School Improvement Planning process and documents required by AYP determinations
- Education Assistance Programs (EAP) to support efforts with students in need of additional support - Tutoring
- Statewide Benchmark Assessment Implementation
- Statewide Progress Monitoring Initiatives
- Statewide Mathematics Initiatives
- Statewide Reading Initiatives
- Distinguished Educator Supports - Technical assistance supports to schools in Corrective Action
- Governor's Institutes: Professional Development Institutes on Accountability and School Improvement
- Higher Education
- Pennsylvania Inspired Leadership (PIL)– a leadership program for school principals

Appendix C - References

Hill, R., Gong, B., Marion, S., DePascale, C., Dunn, J., and Simpson, M.A. (2005, November).

Using value tables to explicitly value student growth. *Paper presented at the Conference on Longitudinal Modeling of Student Achievement, University of Maryland*.

Human Resources Research Organization. (2004) PSSA Issues and Recommendations. *Prepared for the Central Susquehanna Intermediate Unit, Lewisburg, PA.*

Lissitz, Robert W. & Huynh Huynh (2003). Vertical equating for state assessments: issues and solutions in determination of adequate yearly progress and school accountability. *Practical Assessment, Research & Evaluation*, 8(10). Available online: http://ericae.net/pare/getvn.asp?v=8&n=10.

McCaffrey, D., Lockwood, J.R., Koretz, D., Hamilton, L. (2003) *Evaluating Value-Added Models for Teacher Accountability*. Rand Corporation. Santa Monica, CA.

Millman, J. (Ed.) (2004) *Grading Teachers, Grading Schools: Is Student Achievement a Valid Evaluation Measure.* Corwin Publishing, Inc. Thousand Oaks, CA.

Spellings, M. (2005). Letter to Chief State School Officers.

The Council of Chief State School Officers. (2005). *Policymakers' Guide to Growth Models for School Accountability: How Do Accountability Models Differ?* A paper commissioned by the CCSSO Accountability Systems and Reporting, Washington, D.C.

Tucker, P., Stronge, J. (2005) *Linking Teacher Evaluation and Student Learning*, Association for Supervision and Curriculum Development, Alexandria, VA.

Value-Added Assessment Special Issue. (2004). *Journal of Educational and Behavioral Statistics.* American Educational Research Association. Washington, D.C.

Wright, S.P., Sanders, W.L., and Rivers, J.C. (2005), "Measurement of Academic Growth of Individual Students toward Variable and Meaningful Academic Standards", in R. W. Lissitz (ed.) *Longitudinal and Value Added Modeling of Student Performance,* Maple Grove, MN, JAM Press.

Appendix D - Members of the Pennsylvania Value-Added Assessment System Statewide Work Group

PVAAS Statewide Workgroup

First Name	Last Name	Organization
Bonita	Allen	Title I State Parent Advisory Council
Caroline	Allen	State PTA
Joseph	Bard	PA Association of Rural and Small Schools (PARSS)
James	Barker	Erie City SD
Rob	Brown	Pennsylvania Department of Education
Esther	Bush	Urban League of Pittsburgh
Diane	Castelbuono	Pennsylvania Department of Education
Daniel	Collins	PA Association of Elementary and Secondary School Principals (PAESSP)
Robert	Cormany	PA Association of Pupil Services Administrators (PAPSA)
Ronald	Cowell	Education Policy and Leadership Center (EPLC)
Linda	Croushore	Mon Valley Education Consortium
Edward	Donley	State Education Board
Nina	Esposito-Visgitis	PA Federation of Teachers (PAFT)
Dwight	Evans	PA House of Representatives
Thomas	Gentzel	PA School Boards Association (PSBA)
Karl	Girton	State Education Board
James	Goodhart	PA Leauge of Urban Schools
Rodney	Green	Everett Area SD
Paul	Healey	PA Association for Supervision and Curriculum Development (PASCD)
Robert	Hendrickson	College of Education-Penn State University
Stanley	Herman	University of Pittsburgh
Jay	Himes	PA Association of School Business Officials (PASBO)
Linda	Hippert	South Fayette SD
Sharon	Kirk	DuBois Area SD
Michael	McCarthy	PA Business Roundtable
Stephen	Mitchell	Allegheny Conference on Community Development
David	Monk	College of Education-Penn State University
Karen	Murphy	College of Education-Penn State University
Shula	Nedley	Pennsylvania Department of Education
Katherine	Needham	Allegheny Conference on Community Development
Harold	Ohnmeis	Association for Charter Schools
Tim	Potts	PA School Reform Network
Tim	Quinn	Derry Township SD
James J.	Rhoades	PA State Senate
Deb	Rodes	Learning Disability Association
James R.	Roebuck	PA House of Representatives
James	Scott	Lancaster-Lebanon Intermediate Unit 13
Steve	Seliy	Mon Valley Education Consortium
Jim	Shields	PA Association of Intermediate Units (PAIU)
Kevin	Shivers	National Federation of Independent Business
Frederick	Smeigh	Distinguished Educator
Jess M.	Stairs	PA House of Representatives
Janet	Stotland	Education Law Center
Stinson	Stroup	PA Association of School Administrators (PASA)
Louise	Thieme	Parent Education Network
Walter	Vicinelly	Albert Gallatin SD
Edward	Vollbrecht	Pennsylvania Department of Education
Fran	Warkomski	Pennsylvania Training and Technical Assistance Network
Gerald	Zahorchak	Pennsylvania Department of Education
Harris	Zwerling	PA State Education Association (PSEA)

Appendix E - Definitions of AYP Methods

<table>
<tr><td>Method</td><td>AYP is met if</td></tr>
<tr><td>Status Measure
Performance Target
Current Year</td><td>Students meet or exceed the following percentages of proficient students:
<table>
<tr><th>Year</th><th>2002-2004</th><th>2005-2007</th><th>2008-2010</th><th>2011</th><th>2012</th><th>2013</th><th>2014</th></tr>
<tr><td>Percent Proficient in Reading</td><td>45</td><td>54</td><td>63</td><td>72</td><td>81</td><td>91</td><td>100</td></tr>
<tr><td>Percent Proficient in Math</td><td>35</td><td>45</td><td>56</td><td>67</td><td>78</td><td>89</td><td>100</td></tr>
</table></td></tr>
<tr><td>Status Measure
Performance Target
Two-year Average</td><td>Students meet or exceed the above listed percentages using a two-year average of student performance.</td></tr>
<tr><td>Status Measure
Performance Target
Confidence Interval</td><td>The 75% confidence interval of the percentage of proficient students contains the percentages listed in the table.</td></tr>
<tr><td>Status Measures
Performance Targets
Two-year Average
Confidence Intervals</td><td>The confidence interval of the percentage of proficient students from the two-year average contains the percentages listed in the table.</td></tr>
<tr><td>Improvement Measures
Safe Harbor</td><td>The percentage of non-proficient students in the current year is a minimum of 10% less than the percentage of non-proficient students in the previous year.</td></tr>
<tr><td>Improvement Measures
Safe Harbor
Confidence Intervals</td><td>The confidence interval of the percentage of non-proficient students in the current year contains values that represent a reduction in the non-proficient students by a minimum of 10% from the percentage in the previous year.</td></tr>
<tr><td>Improvement Measure
Pennsylvania Performance Index (PPI)</td><td>Meeting PPI targets. PPI is a continuous improvement measure that detects, acknowledges, encourages, and rewards changes across the full range and continuum of academic achievement – not limited solely to the proficient level.</td></tr>
</table>

Appendix F - Projection Methodology

From Wright, Sanders, and Rivers (2005, "Measurement of Academic Growth of Individual Students toward Variable and Meaningful Academic Standards", in R. W. Lissitz (ed.) *Longitudinal and Value Added Modeling of Student Performance,* Maple Grove, MN, JAM Press).

The projection methodology estimates an individual student's academic achievement level at some point in the future under the assumption that this student will have an average schooling experience in the future. The basic methodology is simply to use a student's past scores to predict ("project") some future score. At first glance, the model used to obtain the projections appears to be no more complex than "ordinary multiple regression," the basic formula being:

$$\text{Projected\_Score} = M_Y + b_1(X_1 - M_1) + b_2(X_2 - M_2) + \ldots = M_Y + x_i\, b$$

where M_Y, M_1, etc. are estimated mean scores for the response variable (Y) and the predictor variables (Xs). However, several circumstances cause this to be other than a straightforward regression problem.

Not every student will have the same set of predictors; that is, there is a substantial amount of "missing data." The data are hierarchical: students are nested within classrooms, schools, and districts, and the regression coefficients need to be calculated in such a way as to properly reflect this. The mean scores that are substituted into the regression equation also must be chosen to reflect the interpretation that will be given to the projections.

As noted above, a projection is the score that a student would be expected to make assuming that the student has the average schooling experience in the future. The means should therefore be those of an average school within the population of schools of interest. Also, given this interpretation, the nesting needs to be carried only to the school level (students within schools); it is not necessary to carry it to the classroom level.
The missing data problem can be solved by finding the covariance matrix of all the

predictors plus the response, call it C, with submatrices C_{XX}, C_{XY} (and C_{YX} = C_{XY}), and C_{YY} –1. The regression coefficients (slopes) can then be obtained as b = C_{XX} C_{XY}. For any given student, one can use the subset of C corresponding to that student's set of scores to obtain the regression coefficients for projecting that student's Y value. Because of the hierarchical nature of the data (the second problem), the covariance matrix C must be a pooled-within-school covariance matrix. We obtain this matrix by maximum likelihood estimation using an EM algorithm (to handle missing values) applied to school-mean centered data. Means for an "average school" are obtained by calculating school-mean scores and averaging them over schools. For brevity, we refer to the elements of C, along with the vector of estimated means, as the "projection parameters." Generally, we obtain the projection parameters using the most recent year's data. That is, we use students who have a Y value in the most recent year and X values

from earlier years to get the projection parameters. Projections are then obtained by applying these parameters to students who have X values in the current year (and earlier years) but no Y value.

This methodology does not require vertically linked data nor does it need to assume a linear growth function (or any other specific growth function). Instead, what are required are good predictors of the response variable. The predictors need not be on the same scale with the response or with one another. Potentially, they could be test scores from different vendors and even in different subjects from the response. This gives the methodology considerable flexibility.

Appendix G - Match Rates for 2006 Grades 4 and 6

	4th Grade Students Tested 2006 (after Cleaning)	4th Grade Students 2006 Matched to 3rd Grade 2005	4th Grade Percent Matched across One Year	6th Grade Students Tested 2006 (after Cleaning)	6th Grade 2006 Matched to 5th Grade 2005	6th Grade Percent Matched across One Year
Overall	129,783	116,448	90	127,085	127,890	92
American Indian	172	136	79	171	135	79
Asian/Pacific Islander	3,476	2,975	86	3,406	2,963	87
Black (Non-Hispanic)	20,141	17,455	87	22,342	20,218	91
Hispanic	8,779	7,294	83	9,041	7,618	85
Multi-Racial	906	739	82	681	549	81
White (Non-Hispanic)	95,963	87,770	92	102,037	95,484	94
LEP	1,548	1,061	69	1,453	998	69
Students with Disabilities	19,898	17,859	90	21,094	19,473	92
Economically Disadvantaged	47,341	41,977	89	49,387	45,319	92